Improving Forecast Quality in Practice

Robert Fildes and Fotios Petropoulos

Preview Robert and Fotios discuss the results of their important new survey on the key elements for improving the quality of the forecasting function. A lot more is involved here than software and statistical methodology; it's also about removing organizational impediments, developing appropriate performance benchmarks and motivational incentives, and improving data reliability and flow within the organization.

Key Points

- To improve forecast quality, you must refine all aspects of the forecasting process: the organizational constraints, the flow of information, the forecasting software in use, company resources and techniques employed, and the way you monitor and evaluate forecast accuracy.
- Before attempting to change the forecast function, you should first audit current practices and identify the primary uses as well as the decision-relevant time horizons.
- Survey results show that the primary problem faced in improving forecasting quality is the availability and consistency of internal data. Also of great importance is focused and professional forecasting training as well as the adoption of tested and customisable forecasting systems.
- Improving forecast quality has many aspects. Surveys and case studies provide the evidence that these interact and each one needs to be addressed to reach this improvement target: this means tackling the important areas as an integrated whole.

INTRODUCTION: THE QUALITY OF FORECASTING ACTIVITIES

Foresight readers know there's more to improving organizational forecasting than just developing better statistical algorithms. If you wish to see genuine improvements, you must refine the forecasting process. But exactly how does an organization go about such a formidable if simply stated task?

To gain insight into this question, we began by conducting a survey of the forecasting literature. We found that the topic of *forecast quality improvement* remains a barren field. One of the current coauthors had earlier examined how organizations can improve their forecasting process through a detailed case study of forecasters in 10 divisions of a large multinational corporation (R. Fildes & Hastings, 1994). We will turn to these results later in this paper.

A decade later, Mark Moon, Tom Mentzer, and Carlo Smith (2003) designed a research program into forecasting practice, the results suggesting how organizations should audit their forecasting function. The audit theme was further examined in *Foresight* by Alec Finney and Martin Joseph (2011). These studies all point to the importance of first understanding and evaluating current organizational practice and then examining the gap with best practice.

The second element in our study was a survey of practicing forecasters, asking them about the dimensions of forecasting quality they thought were particularly important in bridging this gap.

Interestingly, more than a few of the insights provided from the earlier studies still hold these many years later.

AUDITING THE FORECASTING FUNCTION

Before an organization can improve the quality of its forecasting, it needs to perform an audit to understand where it currently stands, how its clients perceive forecasting performance, and how its current practice (status quo) stands against appropriate performance benchmarks. This audit should establish a route toward better forecasting and, through an iterative process that identifies how the current forecasting function might change, how to move closer to the "best practice" ideal.

First and most fundamentally, the audit must ask a few essential questions.. Who needs the forecasts? For what purpose? How far into the future do the forecasts need to look (the forecast horizon)? And then there's the question of what added value accrues to the organization by improving the forecast process. After all, without any actual value to improvements, there is really nothing to worry about apart from one's job security. Unfortunately, as readers of this journal are well aware, questions of adding value to forecasting rarely have straightforward answers.

Most forecasts, even if we focus on demand forecasts alone, fulfil a variety of organizational functions – from operations to finance, sales, and marketing. Time horizons vary, too – from short-term operations to midterm budgeting and long-term strategic planning. In their recent *Foresight* article, Petropoulos & Kourentzes (2014) discuss how to combine information from different views of the same data (frequencies) to derive a reconciled set of forecasts suitable for decision making in any functional area.

An exclusive focus, however, on improving the accuracy of the *operational* forecasts may lead to deterioration in the accuracy of forecasts for long-term strategy. That could be the result of selecting methods that perform best only in the short-term or even simply not taking into consideration information that is relevant to long-term planning. So, linking the decision-relevant forecast horizon with performance evaluation is critical for improving overall performance.

Forecast Accuracy Metrics

Foresight has brought forward many discussions on forecasting accuracy metrics as well as a special publication devoted to this topic (*Forecast Accuracy Measurement: Pitfalls to Avoid and Practices to Adopt,* 2012); we will not revisit them in detail here. However, two key conclusions emerge:

1. The accuracy metrics used need to be carefully thought out so that they link to the decisions faced within a specific time horizon.
2. Standard statistical measures won't work with intermittent demands and the corresponding inventory decisions.

Accuracy needs to be measured by taking explicit account of the time horizon faced by the forecast users. For call centres, half-hourly forecasts are often used; for retail operations, it could be within a day; for manufacturing with frozen lead times, forecast horizons are measured in months. For the macroeconomist, horizons range from a "nowcast" (predicting the country's current state) to the next few years ahead, while for the climatologist there is probably no use in mixing up a weather forecast with decadal forecasts of global warming.

While most organizations claim to measure their forecasting accuracy, many of them fail to do so, even when the target is operational, never mind strategic. Key mistakes are:

- The failure to focus on the important forecast horizons. As a result, forecasts are averaged across time horizons, giving a rosy picture of the "achieved" accuracy or unbiasedness.
- Averaging over "apples and oranges." For example, errors of an important product are averaged with errors on unimportant products, completely disregarding product classification strategies.

Benchmarks

To evaluate the current forecasting practice within an organization, forecast accuracy needs to be compared to a suitable benchmark. But what should that benchmark be?

Many organizations tend to evaluate and rank their performance based on surveys of industry practice. This is not a recommended strategy, as forcefully argued by Stephan Kolassa (2008). Every company faces its own problems, and such surveys are almost never representative of them. Moreover, the sampling is usually too small and unscientific to infer statistically valid insights. The responses themselves are not verified, and one organization's customers are not like another's. Most importantly, the philosophy under which each company operates and acts is unique, and this is also reflected by the forecasting function.

What has to be done instead is:

- an off-line comparison using the organization's own data; this could be simply accomplished by withholding the most recent data and producing forecasts based on previous historical information.
- including simple established forecasting methods, but take care to match the method to the data; for example, including seasonality where appropriate
- Implementing top-quality, validated, state-of-the-art software suitable for the needs of the organization; this would guarantee that recent methodological developments will be integrated, while offering the appropriate level of support.

For examples of benchmarking, see the success stories of the Lancaster Centre for Forecasting (http://www.lancaster.ac.uk/lums/forecasting/material/).

The forecast error of current organizational practice needs to be measured relatively to these best-practice benchmarks, as demonstrated by Steve Morlidge (2014) in a recent *Foresight* issue.

So far in the audit, we've highlighted:

- Purpose and value of forecast improvement, the uses of the forecast, and the forecast horizon
- The evaluation of current accuracy/performance and the measurement of forecast error
- Establishing a suitable benchmark

This still leaves additional questions. How are the forecasts produced and with what resources? In particular, what is the information available for use in producing the forecasts? Is the information reliable or even relevant in capturing the true demand (Gilliland, 2010)? Does the utilization of soft information through judgment add value? Are the assumptions underlying the forecasts (such as the promotional plans or future interest rates) open to challenge?

Lastly, an audit must consider the resources that can be mustered: the people, the software, and – not least – the available data. What techniques are used, do they have known flaws, and could they potentially be improved? Are they simple and transparent enough for those in the organization to implement and justify? Does relevant information flow smoothly across the organization and from the outside?

PERSPECTIVES ON IMPROVING FORECASTING QUALITY

From our discussion of the principles behind a forecast audit, the potential problem areas that get in the way of improving quality fall under five headings.

1. Organizational constraints and the flow of information
2. The forecasting software and techniques
3. Limited resources
4. Combining statistical methods with managerial judgment
5. The monitoring and evaluation of the accuracy and value of the forecasting activity

So how can we bridge the performance gap identified through the organizational audit? In the current coauthor's early study of a multinational company and its forecasting functions (Fildes & Hastings, 1994), managers were asked what aspects of their forecasting job were of greatest importance in leading to accuracy improvement. They identified the priorities presented in **Table 1**.

Table 1. Priorities in Improving the Forecasting Accuracy (based on Fildes & Hastings, 1994)

Activity	Respondents scoring the activity as important
Developing consistent data	83%
Increased software support	70%
Improved techniques	66%
Improved availability of easy-to access databases	61%
Improved communication with users	35%

RESULTS FROM THE NEW SURVEY

This early study and, more recently, that by Moon and colleagues (2003) provided the evidence base upon which we designed a questionnaire for demand planners and forecasters. Our primary aim was to explore perceptions of where quality improvements could best be found. To this target, we identified a list of potential problem areas that get in the way of improving the quality of organizational forecasting. This list included 17 problem areas, which could be divided into the five aforementioned headings. We asked the 47 participants in our survey to rate (based on their professional experience) the importance of the issues in improving the quality of the forecasting process. **Figure 1** presents the four most important problem areas, as well as the areas that are not seen as quite so important.

Insert Figure 1 here

It is interesting that even 20 years after the Fildes & Hastings survey, the availability of internal data is still cited as the most important factor in improving the forecast quality. In fact, once again 80% of the responders regard this problem as important or very important. Evaluation (measuring accuracy), organizational issues (motivation to produce better-quality forecasts) and the effective use of judgment (quality of judgmental adjustments made by the forecasting team) complete the top four in terms of importance problem areas.

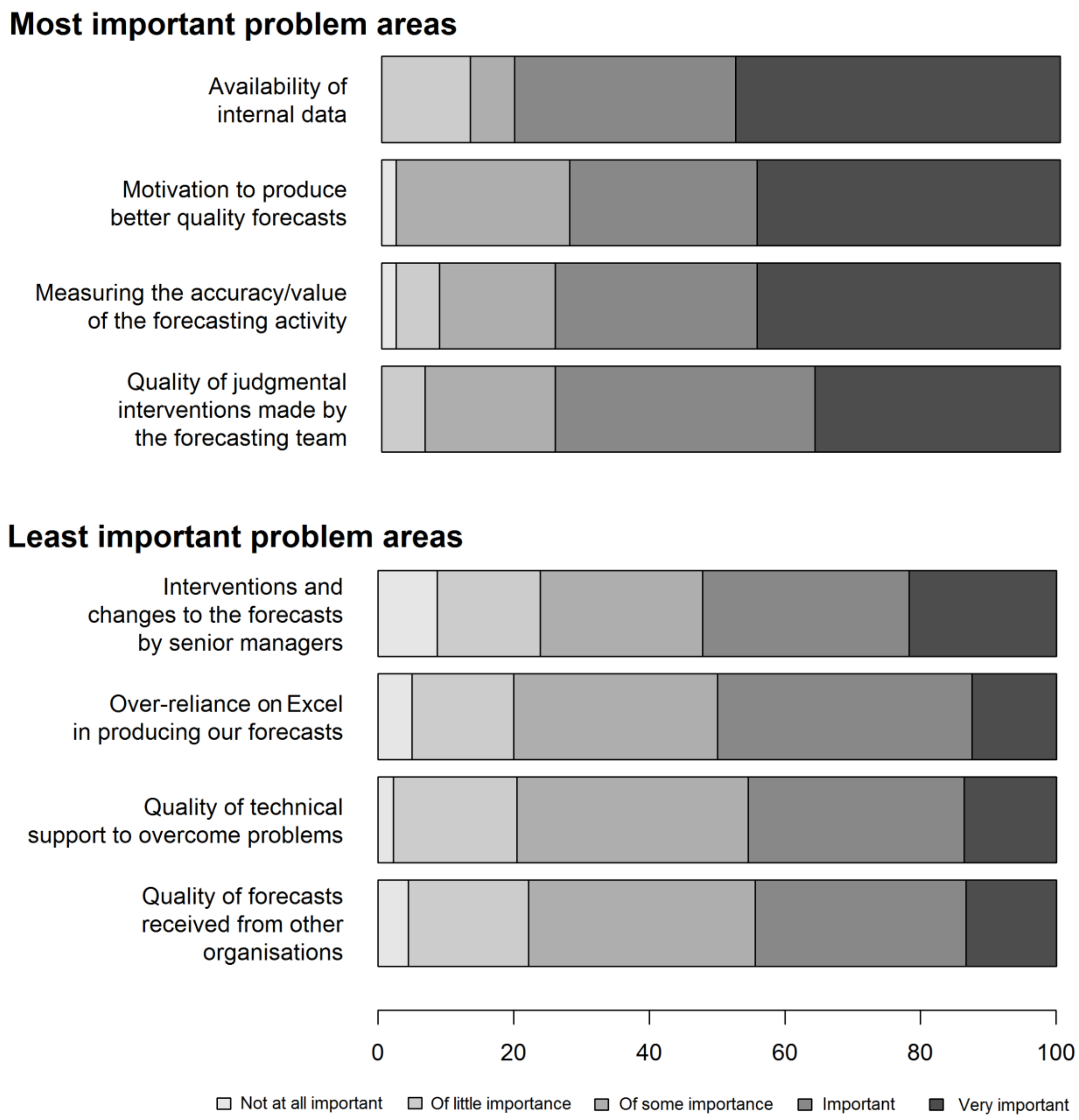

Figure 1. Most and Least Important Problem Areas in Improving Forecast Quality (results from the 2014 survey)

Bottom of the list are those areas that are showing the most improvement over the past years or that are just not identified as important enough:

- Interventions by senior managers, a quite rosy finding especially if we link this with the overoptimistic forecasts/targets that are imposed by the higher levels of management. See Galbraith and Merrill (1996), who document the damage done.

- Over-reliance on Excel, a welcomed result on a first glance, as it indicates that organizations are investing in specialized forecasting tools. But, still 4 in 5 responders find this remains a problem of some importance at least.
- Quality of technical support. This is one area that can be regarded as a win compared to the survey results of Fildes and Hastings (where 70% of the responders found this as important, as opposed to just 45.5% of the responders in our recent survey). An alternative explanation which accords more with our case study experience is that forecasters no longer expect to draw on support for technical concerns but just rely on the available forecasting software.
- Quality of external forecasts. This raises questions with regard to the extent of information sharing, confirmed also by the fact that 21% of the responders have marked the means of external collaboration as "not relevant". Internal forecasts were regarded as more problematic (and also more common a problem).

Other problem areas in our survey included availability and accessibility of an integrated database (important to 69% of responders, increased compared to the 1994 results), lack of training in the forecasting team (59%), and quality of forecasting software (71%).

Role of Judgment

We also needed to establish how forecasts are typically produced. Previous research had established the importance of judgment (Fildes & Goodwin, 2007), which our survey confirms – see **Table 2**.

Table 2. How Forecasts Are Produced

	Fildes & Goodwin (2007)	This survey (2014)
Judgment alone	25%	15.6%
Statistical methods exclusively	25%	28.7%
An *average* of a statistical forecast and management judgmental forecast(s)	17%	18.5%
A statistical forecast judgmentally *adjusted* by the company forecaster(s)	34%	37.1%
Sample Size	149	47

Other researchers have also delved into the types of statistical methods used, with McCarthy and colleagues (2006) offering the most comprehensive summary. Their key finding was that judgment is extensively used, followed by simple methods such as exponential smoothing. While they offered no evidence on the use of regression models to capture causal features of a forecasting problem, a more recent study (Weller & Crone, 2012) of 200 manufacturing companies demonstrates that these are little used.

We draw two conclusions:

1. Judgment is a key element of any forecasting audit and improvement program.
2. There is the potential to improve accuracy through the use of better techniques.

The Forecasters' Targets

But what are the principal targets of forecasters when carrying out their jobs? In our survey, almost 9 in 10 responders identified *accuracy* as the most important objective. Other principal objectives include timeliness, stock availability, and stability of the forecasts. McCarthy and colleagues (2006) had a longer list, with credibility ranked as equally important. Undoubtedly, the focus of any forecasting activity must be forecast accuracy and the resultant credibility that goes with achieving a strong track record.

With accuracy in mind, the forecaster's job is a mixture of activities that include data management, statistical analysis, judgment calls (in order to take into account special factors), and collaboration with colleagues. We use the problem areas identified earlier to analyse the survey evidence on how forecasters believe quality can be improved.

1. Organizational constraints and the flow of information

How do the forecasters collect the information? Is the information (internal or external) easily available? Is information shared across different departments of organization or even across organizations and by what means? What are the organizational constraints imposed by the higher levels of management?

- Respondents claimed their primary sources of information were internal to the company: from sales (79%), marketing (60%), and, to a lesser extent, finance (38%), production/operations (36%), and marketing research (23%).
- In our survey, the availability of internal data still is at the top of the list of areas for improvement, the same position it occupied 20 years previously in the initial survey of Fildes and Hastings (1994). Anecdotes confirm that data are often not delivered on time with quality gaps, or assumptions made (e.g. no out-of-stock) that are not seen as valid.
- External data was not regarded as important as internal, while the quality of externally provided forecasts was identified as even less important.
- Internal resources were used and shared more often than external ones, with about 1 in 5 respondents not using external information sources at all, a finding collaborated by Weller and Crone (2012).
- A dedicated forecasting support system that would provide a structured collaboration and information sharing, either internal or external, still has limited diffusion. Less than 2 in 5 responders use dedicated software as means of internal collaboration, a ratio that drops to less than 1 in 6 for external collaboration purposes. We believe that a specialized forecasting tool that mirrors the S&OP process would boost successful and structured communication, overcoming the barriers of timeliness and incompleteness.
- The majority of the respondents rely on e-mail communications (89%) and meetings (79%) as the main means of internal information sharing and collaboration.
- The good news, compared with earlier surveys, is that senior management support and involvement in intervening with the forecasts are now seen as less important constraints on improving the forecast quality. Motivating the forecasting team remains a requirement for success relying on senior management to recognize and reward forecasting quality.

2. The forecasting software and techniques

Database-related problems still remain a primary concern, even after all these years of hype as to data warehousing and, more recently, "big data". Despite its relatively low ranking, over-reliance on Excel-based solutions is still an important problem for 50% of the responders. The same issue was also identified in an earlier survey by Sanders and Manrodt (2003). One of the organizations in our study was fully reliant on Excel, with data being manually copied from sheet to sheet with little validity testing. In another case, forecasters were confronting size constraints imposed by Excel. As a result, little of the forecasting process was automated.

Our surveys of commercial forecasting software also suggest that the statistical techniques embodied in forecasting support systems have serious limitations (Fildes & Goodwin, 2012). The sub-optimal of even fixed values for model parameters, the inability to take into account multiple seasonal patterns and a reliance on limited data (three years still remains common), lack of features to include explanatory variables (through regression) and the lack of a flexible graphical interface are just some of them. In addition, such systems do not focus adequately on specific user requirements and the unique features of each company's forecasting needs, such as the inclusion of relevant internal or external information or appropriately forecasting the demand of a new product in the early stages of its life-cycle.

3. Limited resources

Lack of forecasting training was perceived as the key resource problem. Without executive, tailored-made training to the needs of each organization, forecasters are liable to fall back on the routines they are most familiar with.

The day-to-day workload also interferes with an improvement programme (1 in 2 found this was an important or very important problem area), suggesting that data management and forecasting have not been successfully routinized, and too much time is spent making judgment calls. In addition, forecasters with too much to do produce poorer forecasts!

4. Combining statistical methods with managerial judgment

It is no surprise to discover that judgment is a key element in the forecasting process. However, forecasters identified the quality of judgmental adjustments made by the forecasting team as an important problem area. They are well aware that they intervene too often, despite basing their adjustments on often inadequate information. While the S&OP process aims to ensure all the information is available when adjustments are made, one of its effects is to make interventions more common and the balance between the statistical forecasts and judgment tipping too far to the latter.

Previous studies suggest that there are ways to improve the quality of such interventions (Fildes & Goodwin, 2007). However, the limitations of current software, which include the lack of provision of guidance and performance feedback for the statistical, judgmental and final forecast separately, are not conducive to supporting better forecasting (Fildes & Goodwin, 2012). What is important is that the value added derived from the S&OP process is measured.

5. Performance evaluation and monitoring

Effectively measuring accuracy remains a problem. Almost 3 in 4 responders identify this as important or very important area. Experience shows that measuring the quality of the forecasting activity in an organization set-up is too rarely done well, if done at all. Metrics need to be chosen carefully to take into account such features as seasonality and to exclude outliers. Some organizations face more difficult forecasting probems (and difficult accounts) than others and should not be judged by the same standards – individual benchmarks may sometimes be needed. Also, the situations where costs of making an error are not symmetric should be considered: this is where it is important to distinguish between the forecasts and the decisions, for example to overstock. If the two are not separated out, it is difficult to improve either the forecasting or the consequential decisions (for example, on service or stocks).

Effective monitoring of the forecasting function is a rarity as well. This failure arises in part because of the lack of adequate measures but also because of the daily workload of the forecaster when meeting the day-to-day demands of producing the regular forecasts. Deficiencies in monitoring stem from software limitations, such as the lack of error bounds that signal unexpected errors. This requires an ABC/XYZ classification, which refers to the proper identification of not only the important SKUs (e.g. in terms of total revenue or profit) but also those that are hardest to forecast. This is yet another requirement for an effectively implemented forecasting support system. In short, despite the technical (and some would say boring) aspects of the problem, measured accuracy is a KPI and is therefore a priority in any forecasting quality improvement programme.

CONCLUSIONS

Improving forecast quality has many dimensions. Surveys and case studies provide the evidence that we need a holistic approach to reach this target, which means tackling the important areas as an integrated whole. Respondents in this most recent survey tended to see the entire list of 17 potential problem areas as (at least) of some importance. So where do we, the authors, see the major improvements arising based on the evidence?

1. Reappraisal of the focus of the forecasting activity; in particular, lead times and level of aggregation suitable for the organization's forecast users
2. Improved and expanded information and an integrated database
3. Benchmarking existing techniques against "best" practice
4. Developing forecasting support systems to manage effective inclusion of judgment.
5. Effective organizational links so that key pieces of information are shared in a timely fashion
6. Trained, motivated, and better-resourced managers.

In short, better data, better software, a wider range of reliable information processed by validated statistical methods, and applied by trained forecasters – now there's a vision for all to work toward!

References

Boylan, J., & Syntetos, A. (2006). Accuracy and accuracy-implication metrics for intermittent demand. *Foresight: The International Journal of Applied Forecasting, 4*, 39-42.

Fildes, R., & Goodwin, P. (2007). Against your better judgment? How organizations can improve their use of management judgment in forecasting. *Interfaces, 37*, 570-576.

Fildes, R., & Goodwin, P. (2007). Good and bad judgment in forecasting: lessons from four companies. *Foresight: The International Journal of Applied Forecasting*, 5-10.

Fildes, R., & Goodwin, P. (2012). Guiding Principles for the Forecasting Support System. *Foresight: The International Journal of Applied Forecasting*, 10-15.

Fildes, R., & Hastings, R. (1994). The organization and improvement of market forecasting. *Journal of the Operational Research Society, 45*, 1-16.

Finney, A., & Joseph, M. (2011). Getting Your Forecasting and Planning Fundamentals Right. *Foresight: The International Journal of Applied Forecasting*, 29-36.

Galbraith, C. S., & Merrill, G. B. (1996). The politics of forecasting: Managing the truth. *California Management Review, 38*, 29-43.

Gilliland, M. (2010). Defining “demand” for demand forecasting. *Foresight: The International Journal of Applied Forecasting, 18*, 4-8.
Hyndman, R. J. (2006). Another look at forecast-accuracy metrics for intermittent demand. *Foresight: The International Journal of Applied Forecasting, 4*, 43-46.
Kolassa, S. (2008). Can we obtain valid benchmarks from published surveys of forecast accuracy. *Foresight, 11*, 6-14.
McCarthy, T. M., Davis, D. F., Golicic, S. L., & Mentzer, J. T. (2006). The evolution of sales forecasting management: A 20-year longitudinal study of forecasting practices. *Journal of Forecasting, 25*, 303-324.
Moon, M. A. (2006). Breaking down barriers to forecast process improvement. *Foresight: The International Journal of Applied Forecasting*, 26-30.
Moon, M. A., Mentzer, J. T., & Smith, C. D. (2003). Conducting a sales forecasting audit. *International Journal of Forecasting, 19*, 5-25.
Morlidge, S. (2014). Using Relative Error Metrics to Improve Forecast Quality in the Supply Chain. *Foresight: The International Journal of Applied Forecasting*, 39-46.
Petropoulos F. & Kourentzes N. (2014) “Improving forecasting via multiple temporal aggregation”, *Foresight: The International Journal of Applied Forecasting*, *34*, 12-17.
Sanders, N. R., & Manrodt, K. B. (2003). Forecasting software in practice: Use, satisfaction, and performance. *Interfaces, 33*, 90-93.
Weller, M., & Crone, S. F. (2012). Supply Chain Forecasting: Best Practices & Benchmarking Study. Lancaster University. http://www.lancaster.ac.uk/lums/forecasting/material/

Authors

Robert Fildes
Lancaster Centre for Forecasting, Department of Management Science, Lancaster University Management School, Lancaster University
r.fildes@lancaster.ac.uk

Robert is Director of the Lancaster Centre for Forecasting and is a past President of the International Institute of Forecasters. He has carried out extensive studies of organizational forecasting practices and the design of forecasting support systems.

Fotios Petropoulos
Logistics & Operations Management Section, Cardiff Business School, Cardiff University
petropoulosf@cardiff.ac.uk

Fotios is Lecturer (Assistant Professor) at Cardiff Business School of Cardiff University. Before that, he was a member of the Lancaster Centre for Forecasting of Lancaster University and the Forecasting & Strategy Unit of the National Technical University of Athens. Fotios is engaged in research on improving forecasting processes.